Achievable Rates with Imperfect Transmitter Side Information Using a Broadcast Transmission Strategy

Avi Steiner and Shlomo Shamai (Shitz)

Abstract

We investigate the performance of the broadcast approach for various fading distributions, which correspond to different models of partial transmit channel state information (CSI). The first model considered is the quantized limited feedback. In this model, the receiver can send as feedback only a finite number of bits describing the fading gain. We derive the optimal power allocation for the broadcast approach for the quantized feedback model. For a Rayleigh fading channel, numerical results here show that if the feedback word can be longer than one bit, the broadcasting gain becomes negligible, due to diminished channel uncertainty. The second partial transmit CSI model is a stochastic Gaussian model with mean and variance information, which is commonly used for modeling the channel estimation error. In a single-input single-output (SISO) channel, this model also corresponds to the Ricean fading distribution, for which we derive maximal achievable broadcasting rates. We further consider a multiple-input single-output (MISO) channel, and derive the optimal power allocation strategy in a broadcast approach. Numerical results here show that uniform power allocation is preferable over beamforming power allocation in the region where broadcasting gain over single level coding is non-negligible.

Index Terms

Broadcast approach, code layering, channel state information, quantized feedback.

I. INTRODUCTION

Transmitter channel state information (CSI) can be useful for increasing achievable rates of a wireless communication system. However, in a rapidly changing environment, transmit CSI (TCSI) may be hard to acquire, as it requires availability of an uplink channel from receiver to transmitter, and even then causal TCSI can be inaccurate in a rapidly varying channel, due to the delay between the channel estimation and the transmission opportunities. In this paper, two models of partial TCSI are considered, namely, a quantized feedback model, and a stochastic Gaussian model.

A quantized feedback model can be found when there is a limited uplink bandwidth. For example, in the UMTS standard there are two bits available for feedback. In [1], the authors study the efficiency of using feedback bits, which determine the quadrant between relative phase of the transmit channels. Power adaptation for applications with long-term power constraint with a quantized feedback is studied in [2], based on the classical time-domain water-pouring approach with perfect TCSI [3]. A generally quantized feedback is studied in [4], in the context of multi-level coding. The authors in [4] use a finite code layering and optimize the quantized information as to maximize the throughput. The broadcast approach [5] is intended for maximizing the expected throughput over fading channels, and for channels with one degree of freedom it is also the ultimate achievable upper bound [6]. In our work, we consider the same channel model as in [4], and derive the maximal throughput achievable adhering to the continuous broadcast approach. The broadcast approach for a single-user facilitates reliable transmission rates adapted to the actual channel conditions, in presence of channel uncertainty at the transmitter [5], [7]. The single-user broadcast approach hinges on the broadcast channel, which was first explored by Cover [8]. In a classical broadcast channel every transmission is composed of multiple messages, each directed to another receiver, where different channel conditions may characterize each receiver, which is reflected in their received SNR. In the broadcast approach [5] considered here, there is only one transmitter and one receiver, which communicate over a fading channel. For channels with one degree of freedom, the multi-level codeword can be (optimally) successively decoded up to highest layer associated with an instantaneous channel realization. The rest of the layers cannot be decoded, and thus will be in outage. Numerical results demonstrate here that with more than one feedback bit the broadcasting gains relative to the outage approach become negligible, as the level of CSI uncertainty diminishes rapidly. The feedback scalar quantization is motivated by [9], which showed that for causal source coding, a scalar quantizer followed by entropy compression is optimal.

The stochastic Gaussian model includes a mean and variance side information of the fading coefficients at the transmitter. This model is useful, for example, in a time-division duplex (TDD) system, where the transmitter can perform channel estimation of inputs from the uplink channel, and use it for transmission. Often, the channel estimation acquired independently by the transmitter may suffer from channel estimation inaccuracy due to RF chain impairments, which limit channel estimation reciprocity. Furthermore, causality requires acquiring CSI prior to transmission, while the channel may change when actual transmission takes place. These give rise to the practical stochastic Gaussian model with mean and variance. This model was used in [10], where it was shown that in the cases of small CSI uncertainty simple beamforming in the direction of the mean channel is quite close to the optimum strategy. Optimal transmission strategies for this model, in terms of outage probability, were studied in [11], [12]. It was shown in [13], that the beamforming is the optimal strategy maximizing the compound capacity of a rank one Ricean channel. We show here that for the broadcast approach in a multi input single output (MISO) channel the optimal power allocation is a linear combination of beamforming and uniform power allocation to the null-space of the known channel.

The rest of the paper is organized as follows. The broadcast approach with quantized feedback is discussed in section II, and for the stochastic Gaussian partial TCSI in section III. It is extended to a MISO channel in section IV. Finally, we conclude with section V.

II. Broadcasting with Quantized Feedback

Consider a slowly fading channel, where a limited uplink channel sends only a predetermined number of information bits as feedback. In a channel with one degree of freedom (SISO/SIMO/MISO) this information can be a quantized version of the equivalent fading gain. It may indicate a fading gain range, as observed by the receiver. Transmission is then optimally adapted to the fading gain feedback via an outage strategy or a broadcast strategy.

Figure 1, illustrates an exemplary communication protocol, which enables transmit side information via a quantized feedback. Every transmission is initiated with a request-to-send (RTS) message. The receiver replies with a clear-to-send (CTS) indication, and adds to the CTS message N bits describing the link quality. Then the transmitter adapts its transmission rate and broadcast strategy to this feedback.

Consider the following SISO channel model,

$$\mathbf{y} = h\mathbf{x} + \mathbf{n} \ , \tag{1}$$

where $\mathbf{y}$ is a received vector of length L, corresponding to the number of channel uses per transmission block. The channel input $\mathbf{x}$ is the transmitted vector of length L, which satisfies the power constraint $E\frac{1}{L}\sum\limits_{i=1}^{L}|x_i|^2 \leq P$, where x_i is the i^{th} transmitted symbol within a block. The additive noise vector is denoted $\mathbf{n}$ and its elements are complex Gaussian i.i.d with zero mean and unit variance, denoted by $\mathcal{CN}(0,1)$. The channel fading coefficient h is assumed to remain fixed during a transmission block. Over multiple transmission blocks h admits the normalized Gaussian distribution $\mathcal{CN}(0,1)$. Partial TCSI information modifies the fading coefficient distribution, with implications on achievable rates.

Throughout this paper, only the short term power constraint P is imposed, that is $E\frac{1}{L}\sum\limits_{i=1}^{L}|x_i|^2 \leq P$ per block. We note here that under a long term power constraint (average power constraint over multiple transmission blocks), higher average throughput may be achievable with a different optimal power allocation, since the quantized feedback indicates a coarse channel quality, and a water-pouring like strategy is required to maximize throughput [2], [3].

Before the analysis of the broadcast approach with partial CSI, we formulate the perfect TCSI upper bound. This is also the ergodic capacity, which is explicitly given for the Rayleigh fading case,

$$C_{erg}^{Rayleigh} = E_\nu \log(1 + P\nu) = e^{1/P} E_1(1/P) \tag{2}$$

where ν is Rayleigh distributed, and $E_1(x) = \int_x^\infty \frac{e^{-t}}{t} dt$ is the exponent integral function.

A. Single Bit of Transmit Side Information

Consider first the case that only 1-bit of side information is available at the transmitter. The feedback indicates whether the actual fading gain is above or below some threshold s_{th}. The average achievable rate is given, in general, by

$$R_{avg} = P_\nu(\nu \leq s_{th}) R_{avg,1} + P_\nu(\nu > s_{th}) R_{avg,2} \tag{3}$$

where ν is the channel fading gain random variable, and s_{th} is the fading gain threshold. The rates $R_{avg,1}$ and $R_{avg,2}$ are average achievable rates for $\nu \leq s_{th}$, and $\nu > s_{th}$, respectively.

1) Outage Approach: In a single level coding scheme (outage approach), there is a single fixed rate code per transmission block. Therefore, in case of severe channel conditions, nothing is decoded, and an outage event occurs. When fading gain is above some level all transmitted information is reliably decoded. The maximal achievable average throughput is known as the outage capacity [14]. For a given threshold selection s_{th}, the rates $R_{avg,1}$, and $R_{avg,2}$ are average achievable rates of single level coding. The achievable rate, for a given feedback ($k = 1, 2$), is

$$R_{avg,k} = (1 - P_{o,k}) R_k \tag{4}$$

where $P_{o,k}$ is the outage probability on a k^{th} fading gain interval, and R_k is the corresponding single level code rate allocated, where $k = 1, 2$ for a single bit feedback. Each of the average rates $R_{avg,k}$ may be optimized independently, for some s_{th}, such that R_{avg} (3) is maximized. The outage probability for $\nu \leq s_{th}$, and a transmission in a fixed rate $R_1 = \log(1 + u_1 P)$, is

$$P_{o,1} = P_\nu(\nu < u_1 | \nu \leq s_{th}) = \frac{1 - e^{-u_1}}{1 - e^{-s_{th}}}. \quad (5)$$

The corresponding maximal average rate is

$$R_{avg,1} = \max_{u_1 \in [0, s_{th}]} (1 - P_{o,1}) \log(1 + u_1 P) = \max_{u_1 \in [0, s_{th}]} \frac{e^{-u_1} - e^{-s_{th}}}{1 - e^{-s_{th}}} \log(1 + u_1 P). \quad (6)$$

Similarly, for a given threshold u_2, the probability of outage for $\nu > s_{th}$ is

$$P_{o,2} = P_\nu(\nu < u_2 | \nu > s_{th}) = e^{-(u_2 - s_{th})}. \quad (7)$$

The maximal average rate is then

$$R_{avg,2} = \max_{u_2 \in (s_{th}, \infty)} (1 - P_{o,2}) \log(1 + u_2 P) = \max_{u_2 \in (s_{th}, \infty)} e^{-(u_2 - s_{th})} \log(1 + u_2 P), \quad (8)$$

where (8) is maximized for

$$u_2^{opt} = \max\left(\frac{P - W(P)}{P \cdot W(P)}, \quad s_{th} \right) \quad (9)$$

where $W(x)$ is the Lambert W-function, also called the omega function. The selection of s_{th} can be numerically optimized, to maximize the total average rate R_{avg} in (3) for the outage approach.

2) Broadcast Approach: The achievable outage average rates may be further enhanced by using a broadcast approach [5] optimized for every channel state. We derive here the continuous broadcast approach for $R_{avg,1}$ and $R_{avg,2}$ in (3). The broadcast approach for the SISO channel was introduced in detail in [7]. For completeness of presentation, we quickly review the principles of the broadcast approach. The concept of the broadcast approach is demonstrated in Figure 2. As may be noticed, in the outage approach transmission is either reliably decoded or completely fails (outage). Whereas in the broadcast approach the number of decoded layers varies w.r.t. the channel fading gain. The undecoded layers induce an outage region. In the continuous broadcast approach, every layer is associated with a fading gain $u = |h|^2$, and is allocated a fractional power $\rho(u)du$. Thus the incremental rate associated with layer u is [7]

$$dR(u) = \log\left(1 + \frac{\rho(u)udu}{1 + I(u)u}\right) = \frac{\rho(u)udu}{1 + I(u)u} \tag{10}$$

where $I(u)$ is the residual interference function, such that $I(0) = P$, and $\rho(u) = -\frac{d}{du}I(u)$ is the power allocation density function. The maximal average rate is expressed as follows

$$R_{bs,avg} = \max_{I(u)} \int_0^\infty du(1 - F_\nu(u))\frac{\rho(u)u}{1 + I(u)u} \tag{11}$$

where $F_\nu(u)$ is the cumulative distribution function (cdf) of the fading gain random variable. It may be shown [7] that the optimal residual interference distribution function is given by

$$I_{opt}(u) = \begin{cases} P & u < u_0 \\ \frac{1 - F_\nu(u) - u \cdot f_\nu(u)}{u^2 f_\nu(u)} & u_0 \le u \le u_1 \\ 0 & u > u_1 \end{cases} \tag{12}$$

where u_0 and u_1 are obtained from the boundary conditions $I_{opt}(u_0) = P$, and $I_{opt}(u_1) = 0$, respectively. From $I(u)$, the power density function is directly obtained by $\rho_{opt}(u) = -I'_{opt}(u)$, which yields the optimal layering power distribution. In order to study the benefits of layering, the broadcast approach is always compared to the single level coding under the same channel uncertainty.

When a single feedback bit is available at the transmitter, two different broadcast approaches are to be used, $R_{avg,1}$ corresponding to the first state described by $\nu \le s_{th}$, and $R_{avg,2}$, corresponding to $\nu > s_{th}$. The average broadcasting rate for $\nu \le s_{th}$, is specified by

$$R_{avg,1} = \int_0^{s_{th}} (1 - F_1(u))\frac{\rho_1(u)u}{1 + I_1(u)u}du \tag{13}$$

where $F_1(u) = \frac{1-e^{-u}}{1-e^{-s_{th}}}$. Similarly, for $R_{avg,2}$ in (3), when $\nu > s_{th}$, the broadcasting average rate is specified by,

$$R_{avg,2} = \int_{s_{th}}^{\infty} (1 - F_2(u)) \frac{\rho_2(u)u}{1 + I_2(u)u} du \tag{14}$$

where $F_2(u) = \frac{e^{-s_{th}} - e^{-u}}{e^{-s_{th}}}$. The maximal achievable rate is specified by the following proposition.

Proposition 2.1: Maximal Achievable rates, in presence of a single bit feedback for TCSI, is given for a Rayleigh fading channel by

$$R_{avg,1,max} = (1-e^{-s_{th}})^{-1}(2E_1(s_0^{(1)})-2E_1(s_1^{(1)})+e^{-s_1^{(1)}}-e^{-s_0^{(1)}}+e^{-s_{th}}(s_1^{(1)}-s_0^{(1)}-2\log\left(\frac{s_1^{(1)}}{s_0^{(1)}}\right))) \tag{15}$$

$$R_{avg,2,max} = \begin{cases} e^{s_{th}}(2Ei(s_0^{(2)}) - 2Ei(s_1^{(2)}) - & \\ (e^{-s_0^{(2)}} - e^{-s_1^{(2)}})) & s_{th} < s_0^{(2)} \\ 2e^{s_{th}}\left(Ei(s_{th}) - Ei(s_1^{(2)})\right) - 1 + & \\ e^{-(s_1^{(2)} - s_{th})} + \log[(1 + s_{th}P)s_{th}] & s_0^{(2)} \le s_{th} \le s_1^{(2)} \\ \log(1 + s_{th}P) & s_{th} > s_1^{(2)} \end{cases} \tag{16}$$

where $s_0^{(k)}$, and $s_1^{(k)}$ are determined by requiring separately in (15)-(16) that $I_k^{opt}(s_0^{(k)}) = P$, and $I_k^{opt}(s_1^{(k)}) = 0$, respectively, and $k = 1, 2$. The expected rate is then given by (3).

Proof: See Appendix A.

B. Multi-level Transmit Side Information

When the receiver is capable of sending as feedback a higher resolution quantized fading gain, the transmitter can perform closer to the ergodic capacity upper bound (2). Consider a K level feedback, which indicates that the fading gain belongs to one of K possible intervals, represented by $N = \lceil \log_2(K) \rceil$ bits. Let the intervals be denoted by

$$\{\mathcal{I}_k\}_{k=1}^{K} = \{[s_{th,0}, s_{th,1}), [s_{th,1}, s_{th,2}), ..., [s_{th,K-1}, s_{th,K})\} \tag{17}$$

where $\mathcal{I}_k = [s_{th,k-1}, s_{th,k})$ with $s_{th,0} = 0$, and $s_{th,K} = \infty$. The average achievable rate given K-level transmitter side information is generally given by

$$R_{avg} = \sum_{k=1}^{K} P_\nu(\nu \in \mathcal{I}_k) \cdot R_{avg,k} \tag{18}$$

where ν is the channel fading gain random variable, $R_{avg,k}$ is the average rate for the k^{th} channel state interval, and $P_\nu(\nu \in \mathcal{I}_k)$ is the probability that the channel state is within an interval $\mathcal{I}_k$.

1) Outage Approach: Consider first the single level coding (outage) approach. For some interval $\mathcal{I}_k$, the rate $R_{avg,k}$ may be optimized for maximizing the overall throughput R_{avg} (18). Let u_k determine the k^{th} fading gain threshold by fixing the k^{th} transmit rate to $R_k = \log(1+u_k P)$. Then the outage probability with side information $\mathcal{I}_k$ is

$$P_{o,k} = P_\nu(\nu < u_k | \nu \in \mathcal{I}_k) = \frac{\exp(-s_{th,k-1}) - \exp(-u_k)}{\exp(-s_{th,k-1}) - \exp(-s_{th,k})} \tag{19}$$

where it is implicitly assumed that $u_k \in \mathcal{I}_k$. The maximal average outage rate is then

$$R_{avg,k} = \max_{u_k \in \mathcal{I}_k} (1 - P_{o,k}) \log(1+u_k P) = \max_{u_k \in \mathcal{I}_k} \frac{\exp(-u_k) - \exp(-s_{th,k})}{\exp(-s_{th,k-1}) - \exp(-s_{th,k})} \log(1+u_k P). \tag{20}$$

2) Broadcast Approach: The optimal continuous broadcast approach can be computed for every fading gain interval $\mathcal{I}_k$. This is summarized in the following proposition.

Proposition 2.2: The k^{th} maximal achievable rates, in presence of a K-level feedback, is given by

$$R_{avg,k,bs} = \int_{s_{th,k-1}}^{s_{th,k}} (1 - F_k(u)) \frac{\rho_k(u) u}{1 + I_k(u) u} du \tag{21}$$

with optimal power allocation,

$$I_k^{opt}(s) = \begin{cases} P & s < s_0^{(k)} \\ \frac{1-e^{-(s_{th,k}-s)}}{s^2} - \frac{1}{s} & s_0^{(k)} \le s \le s_1^{(k)} \\ 0 & s > s_1^{(k)} \end{cases} \tag{22}$$

where $s_1^{(k)} = 1 - W\left(e^{-(s_{th,k}-1)}\right)$, and $s_0^{(k)}$ is determined by $I_k^{opt}(s_0^{(k)}) = P$.

Proof: The derivation of the optimal power distribution $I_k^{opt}(s)$ follows the same guidelines of the proof of Proposition 2.1.

Notice that when $s_0 \leq s_{th,k-1}$, there is no point in power allocation for fading gains smaller than the minimal possible fading gain. Therefore, all power intended for lower layers is unified into the first layer, denoted by $s_{th,k-1}$. Thus the power for $s_{th,k-1}$, when $s_0 \leq s_{th,k-1} < s_1$ is $P(s_{th,k-1}) = P - \frac{1-e^{-(s_{th,k}-s_{th,k-1})}}{s_{th,k-1}^2} + \frac{1}{s_{th,k-1}}$. In case $s_{th,k-1} \geq s_1$, then no layering is required, and single level coding is optimal, in a rate corresponding to $R_{avg,k} = \log(1 + s_{th,k-1}P)$.

A common assumption in literature is that the transmitter knows the fading gain pdf, denoted $f(s)$. Our results can be adapted for any given fading distribution $f(s)$, and thus the thresholds $s_{th,k}$ are optimized once, during system configuration, which means the transmitter does not need any additional feedback to determine $s_{th,k}$.

C. Numerical Results

Numerical results for a SISO Rayleigh fading channel are presented in the following. Figure 3 demonstrates the outage and broadcast approaches maximal achievable rates as function of the threshold s_{th}, for a single bit feedback. The receive SNR is 20 dB in this example. As may be noticed, the broadcasting gain over the outage approach strongly depends on the selection of the feedback threshold. This gain is minimized at the point of optimal selection of the feedback threshold ($s_{th} = 0.7$).

Figure 4 demonstrates the broadcasting gain over outage approach with $K = 2$, and $K = 3$ transmit intervals side information. In high SNRs, and $K = 2$, the broadcasting gain over outage is reduced to ∼1 dB in presence of a single bit feedback, relative to a ∼3.5 dB broadcasting gain without feedback. For $K = 3$, the broadcasting gain practically diminishes. Note also that when $K = 3$ the feedback has to include $N = 2$ bits.

III. Broadcasting with partial CSI adhering to the Gaussian Model

The case of quantized feedback is not always a valid model for imperfect transmitter side information. In a TDD system, for example, the transmitter may perform uplink channel estimation and use it for downlink transmission. However, the channel estimation may be limited due to many reasons, e.g. the channel estimation acquired independently by the transmitter may suffer from inherent inaccuracy due to RF chain impairments, which limit channel reciprocity. Therefore, we consider now a SISO channel, as defined in (1), with a stochastic TCSI model [10], [15]. The transmitter performs broadcasting, in order to maximize the average throughput under this Gaussian channel model. The average throughput is compared with the outage capacity lower bound, and the ergodic capacity upper bounds.

A. SISO Broadcasting over a Ricean Fading Channel

The channel fading coefficient is described by

$$h = \widehat{h}_t + h_r \tag{23}$$

where $\widehat{h}_t$ represents the estimated fading coefficient, and $h_r \sim \mathcal{CN}(0, \sigma^2)$ is the estimation error random variable. The transmitter knows $\widehat{h}_t$ and σ^2. This channel model corresponds also to the line-of sight model, where $\widehat{h}_t$ represents the channel fading coefficient on the line-of-sight path. Clearly, the square norm of the fading random variable $\nu = |h|^2$ is Ricean distributed. Ricean distribution is a particular case of the noncentral Chi-square (χ^2) distribution, with a noncentrality parameter $|\widehat{h}_t|^2$.

The perfect TCSI ergodic capacity upper bound for the this channel model is given by

$$C_{erg}^{Rice} = E_\nu \log(1 + P\nu) \tag{24}$$

where $\nu = |h|^2$, and for the Gaussian channel model (23), we have $h \sim \mathcal{CN}(\widehat{h}_t, \sigma^2)$, and ν has a Ricean distribution, which is a noncentral χ^2 distribution with two degrees of freedom. Its probability density function (pdf) is [16]

$$f_\nu(u) = \frac{1}{\sigma^2} e^{-\frac{|\widehat{h}_t|^2+u}{\sigma^2}} B_0\left(\frac{2|\widehat{h}_t|\sqrt{u}}{\sigma^2}\right), \quad (25)$$

where $B_0(x)$ is the 0^{th} order modified Bessel function of the first kind.

For single level coding, the outage capacity can be expressed as function of the fading cdf,

$$C^{Rice}_{outage} = \max_{\nu_{th}} (1 - F_\nu(\nu_{th})) \cdot \log(1 + P\nu_{th}) \quad (26)$$

where $F_\nu(x)$ is the cdf of ν, and is also an outage probability. The optimizing fading threshold satisfies $\nu_{th,max} = \arg\max_{\nu_{th}} (1 - F_\nu(\nu_{th})) \cdot \log(1 + P\nu_{th})$. Hence an outage event occurs for channel realizations which satisfy $\nu < \nu_{th,max}$, otherwise a rate $R_1 = \log(1 + P\nu_{th,max})$ is reliably decoded. The broadcast approach [7] throughput can be derived in the same lines of the SISO channel, which was introduced in the previous section. The broadcast approach throughput is specified in (11), and its optimizing power distribution is maximized using Eq. (12), where the fading gain distribution used in (11)-(12) is the Ricean distribution given in (25).

B. Numerical Results

We present here numerical results for the maximal average throughput of the broadcast approach $R^{Rice}_{bs,avg}$ (11), outage approach C^{Rice}_{outage} (26), and the ergodic upper bound C^{Rice}_{erg} (24).

Figures 5.(a)-5.(d) demonstrate the potential gains of broadcasting for different values of σ^2. For $\sigma^2 = 1$, the channel estimation is $\widehat{h}_t = 0$ (or alternatively there is no line of sight), and thus ν_s is Rayleigh distributed. In all other cases $\sigma^2 < 1$, and the gains of broadcasting over the outage approach are lower. The smaller σ^2 the lower the broadcasting gain. For $\sigma^2 = 1$, in high SNRs the broadcasting gain over the outage approach reaches ~ 3.5 dB, and for $\sigma^2 = 0.1$, the broadcasting gain is approximately 1.8 dB.

IV. Broadcasting over the MISO Channel with Partial CSI

When transmitter has more than one antenna, it may utilize both phase and amplitude of the TCSI for maximizing throughput, by using beamforming like strategies. This is unlike the case of one transmit antenna, where only the channel magnitude is helpful. We consider here the MISO channel with partial TCSI under the stochastic Gaussian TCSI model. It was shown in [12] that with such TCSI, and an outage approach, the optimal power allocation is a linear combination of beamforming and uniform power distribution over the null-space. We extend this result for the broadcast approach. Consider the following MISO channel,

$$\mathbf{y} = \mathbf{h}_s \mathrm{X} + \mathbf{n}, \tag{27}$$

where $\mathbf{y}$ is a received vector of length L, corresponding to the number of channel uses per transmission block. The channel input X is the transmitted $[N \times L]$ matrix, which satisfies the power constraint $E\frac{1}{L}\sum_{i=1}^{L} \mathbf{x}^{(i)\dagger}\mathbf{x}^{(i)} \leq P$, where $\mathbf{x}^{(i)}$ is the i^{th}, $[N \times 1]$, transmitted column of X. The $[1 \times L]$ additive noise vector is denoted $\mathbf{n}$ and its elements are complex Gaussian i.i.d with zero mean and unit variance, denoted by $\mathcal{CN}(0, 1)$. The fading channel vector $\mathbf{h}_s$ is a $[1 \times N]$ vector, its elements $h_s(i)$ are Ricean distributed as in (23). It is assumed that the $\{h_s(i)\}_{i=1}^{N}$ are independently distributed each with a specific line-of sight parameter $Eh_s(i) = m_i$, and $\mathbf{m}_t$ is the N element vector of the channel mean available to transmitter. The channel uncertainty random variables $h_r(i)$ are i.i.d $h_r(i) \sim \mathcal{N}(0, \sigma^2),\ i = 1, .., N$. The average mutual information is give by

$$I(\mathbf{x}; y|\mathbf{h}_s) = \log(1 + P\mathbf{h}_s Q \mathbf{h}_s^\dagger) \tag{28}$$

where Q is an $[N \times N]$ covariance matrix of the power distribution, such that $\mathrm{tr}(Q) = 1$.

A. Uniform Power Allocation

For uniform power allocation, $Q = \frac{1}{N} I_N$, where I_N is an $N \times N$ identity matrix. The mutual information for this power allocation is

$$I_{uni}(\mathbf{x}; y|\mathbf{h}_s) = \log(1 + \frac{P}{N}\nu_u), \tag{29}$$

where $\nu_u = \sum_{i=1}^{N} s_i$, with $s_i \triangleq |h_s(i)|^2$. The random variable ν_u has a noncentral χ^2 distribution with N degrees of freedom, and noncentrality parameter $m = \sum_{i=0}^{N} |m_i|^2$, and variance σ^2 to each element in the sum. The pdf of ν_u is given by

$$f_{\nu_u}(x) = \frac{1}{2\sigma^2}\left(\frac{x}{m}\right)^{(N-2)/4} e^{-\frac{x+m}{2\sigma^2}} B_{N/2-1}\left(\frac{\sqrt{xm}}{\sigma^2}\right) \tag{30}$$

where $B_n(x)$ is the n^{th} order modified Bessel function of the first kind. Note that the average receive SNR here is $SNR_{uni} = \frac{P}{N}(m + N\sigma^2)$. Given the mutual information expression (29), and the distribution of the equivalent fading ν_u (30), broadcasting and outage rates may be derived straightforwardly, following same guidelines as in the scalar Ricean fading case, Section III. However, the uniform power allocation for a MISO channel in case of TCSI is sub-optimal. It is therefore interesting to characterize the optimal power distribution for MISO broadcasting with partial TCSI.

B. Optimal Power Allocation

We refer here to the optimal power allocation strategy. The known result for the outage approach strategy [12] is derived first, which is then extended to the broadcast approach.

1) ***Outage Approach:*** Consider the power covariance matrix Q in its SVD representation $Q = U\Lambda U^{\dagger}$. Let $\underline{\nu} = U^{\dagger}\mathbf{h}_s$, it was shown in [12] that for minimal outage probability the independent elements of $\underline{\nu}$ satisfy $\nu_1 \sim \mathcal{N}(\sqrt{m}, \sigma^2)$, and $\nu_2, .., \nu_N \sim \mathcal{N}(0, \sigma^2)$, where $\lambda_1 > \lambda_i$ for all $i = 2, .., N$ without loss of generality. The columns of the unitary matrix U are composed of $\left(\frac{\mathbf{m}_t}{\sqrt{m}}, \mathbf{u}_2, ..., \mathbf{u}_N\right)$, where $\{\mathbf{u}_i\}$ $(i = 2, .., N)$ is an orthonormal set of arbitrary vectors, which are also orthonormal to $\frac{\mathbf{m}_t}{\sqrt{m}}$. This means that λ_1 is the power assignment in the beamforming direction. The other singular values represent power allocation in the null-space, to random variables (RV) $\nu_2, .., \nu_N \sim \mathcal{N}(0, \sigma^2)$, since these are i.i.d. Gaussian RVs equal power allocation is optimal here [12], hence $\lambda_2 = \lambda_3 = \cdots = \lambda_N$ is optimal. The mutual information is then

$$I_{opt}(\mathbf{x}; y|\mathbf{h}_s) = \log(1 + P\left(\lambda_1 s_1 + \beta s_u\right)), \tag{31}$$

which results from $\beta \triangleq \lambda_2 = \lambda_3 = \cdots = \lambda_N$, and by requiring $s_u = \sum_{i=2}^{N} s_i$, where s_u is central χ^2 distributed with $N - 1$ degrees of freedom. The parameter s_1 is a noncentral χ^2 RV with variance σ^2 and noncentrality parameter m. By incorporating the power constraint $\sum_{i=1}^{N} \lambda_i = 1$ into (31)

$$I_{opt}(\mathbf{x}; y|\mathbf{h}_s) \quad = \log(1 + P\left(\alpha s_1 + \tfrac{(1-\alpha)}{N-1} s_u\right)) \tag{32}$$

where $\alpha \in [\frac{1}{N}, 1]$, the requirement $\alpha \geq 1/N$ is a direct result of the optimality condition that $\lambda_1 \geq \lambda_i$ for all $i = 2, .., N$. We are interested in the distribution of the random variable $t = \alpha s_1 + \frac{(1-\alpha)}{N-1} s_u$. Since s_1 and s_u are statistically independent, the pdf of t is given by a convolution of the individual pdfs,

$$f_t(t) = \int_0^{t/\alpha} ds f_{s_1}(s) \cdot f_{s_u}\left(\frac{(N-1)}{1-\alpha} t - \frac{(N-1)\alpha}{1-\alpha} s\right) \cdot \frac{(N-1)}{1-\alpha} \tag{33}$$

where the pdfs of s_1 and s_u are as follows

$$f_{s_1}(s) = \qquad \frac{1}{\sigma\sqrt{2\pi s}} e^{-\frac{s+m}{2\sigma^2}} \cosh\left(\frac{\sqrt{sm}}{\sigma^2}\right) \tag{34}$$

$$f_{s_u}(s) = \frac{1}{\sigma^{N-1} 2^{(N-1)/2} \Gamma(\frac{N-1}{2})} s^{(N-1)/2-1} e^{-\frac{s}{2\sigma^2}} \tag{35}$$

The SNR at a single receive antenna for the above power allocation is given by $SNR_{opt} = P \cdot (\alpha(m + \sigma^2) + (1 - \alpha)\sigma^2) = P \cdot (\alpha m + \sigma^2)$. From here the average outage rates can be optimized like in the uniform power allocation case, using the fading gain distribution in (33).

*2) **Broadcast Approach:*** We have seen that for an outage approach the optimal power allocation strategy is to allocate the larger portion of the power in the beamforming direction, and the rest is allocated to the null-space, uniformly among the eigen-modes. We will show here that this strategy is also optimal for the broadcast approach as a per layer power allocation.

Consider a finite level code layering with n layers, and (scalar) power allocation $P_1, P_2, ..., P_n$ to each layer, such that $\sum_{i=1}^{n} P_i = P$. The corresponding rates are denoted by $R_1, R_2, ..., R_n$. In addition, let Q_k be a power covariance matrix for layer k, such that $\text{tr}\{Q_k\} = 1$. Then, the channel realizations satisfying the rate and power allocation constraints may be successfully decoded. That is, the probability of outage for layer k is specified by,

$$P_{o,k} = \text{Pr}\left(\log\left(1 + \frac{P_k \mathbf{h}_s Q_k \mathbf{h}_s^\dagger}{1 + \sum\limits_{j=k+1}^{n} P_j \mathbf{h}_s Q_j \mathbf{h}_s^\dagger}\right) < R_k\right). \quad (36)$$

The following proposition shows that the covariance matrix Q_k for the broadcast approach has the same properties as in the outage case.

Proposition 4.1: The optimal power allocation strategy for the broadcast approach requires covariance matrices $\{Q_k\}_{k=1}^{n}$, with an SVD representation $Q_k = U \Lambda_k U^\dagger$, where the matrices Λ_k and U have the following properties:

1) $\Lambda_k = \text{diag}\{\lambda_{1,k}, ..., \lambda_{N,k}\}$, such that the largest eigenvalue is $\lambda_{1,k}$ ($\lambda_{1,k} > \lambda_{j,k}$, $\forall j,\ 2 \leq j \leq N$), and $\lambda_{2,k} = \lambda_{3,k} = ... = \lambda_{N,k}$.
2) The unitary matrix U_k includes, for any k, the following columns $\left(\frac{\mathbf{m}_t}{\sqrt{m}}, \mathbf{u}_2, ..., \mathbf{u}_N\right)$, where $\{\mathbf{u}_i\}_{i=2}^{N}$ is an orthonormal set of arbitrary vectors, also orthonormal to $\frac{\mathbf{m}_t}{\sqrt{m}}$.

Proof: See Appendix B.

C. Numerical Results

Two examples of uniform power allocation and beamforming only power allocation are presented. The uniform power allocation average rate follows from (29). Beamforming only power allocation is a special case of the optimal power allocation strategy with $\alpha = 1$ in (32). In Figures 6-7, the channel uncertainty is $\sigma^2 = 0.16$, and $\sigma^2 = 0.01$, respectively. In both figures $m = 1$.

Figure 6 demonstrates a rather high CSI uncertainty of $\sigma^2 = 0.16$, uniform power allocation is preferable. The broadcasting gain over the outage approach exceeds ~ 2 dB. Figure 7, shows that for a lower CSI uncertainty such as $\sigma^2 = 0.01$, the outage and broadcast approach nearly coincide, and the performance of beamforming and uniform power allocation is also very close. Moreover, the proximity of performance to the ergodic capacity is also a direct result of low CSI uncertainty. Clearly, by further increasing the certainty in CSI the beamforming power allocation will be superior, and broadcasting will not be beneficial any longer.

V. Conclusion

We have studied several power allocation strategies for the broadcast approach with partial TCSI, with focus on analytical solutions. Nevertheless, the principles set here can be used for different nature of side information motivated by practical considerations to assess the benefits from a multi-layer broadcast approach. The optimal power allocation and maximal achievable broadcasting rate is derived for the case of a quantized fading gain feedback. It is demonstrated via numerical results that for a Rayleigh fading channel with only one feedback bit, the broadcast approach introduces roughly 1 dB gain over the outage approach. This gain diminishes when increasing the resolution of the quantized information beyond one bit.

We have also considered the stochastic Gaussian TCSI model, where the transmitter possesses only the mean and variance of the actual CSI. Optimal power allocation for the SISO channel is derived similarly to the case where no TCSI is available. The MISO channel with the same channel uncertainty model was also considered, and it was shown that optimal power allocation for the broadcast approach is a linear combination of beamforming and uniform power allocation on all eigen-modes of the null-space.

Throughout the paper the power constraint is a short-term power constraint (per transmission block). It is well known that with a long-term power constraint (over multiple blocks), a water-filling like strategy for transmitter power control achieves the outage capacity. It may be interesting to characterize the achievable rates with a broadcast approach, when power control is allowed, as this is still an open problem, and may be considered in future work.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge constructive discussions with Ami Wiesel, and wish to thank the associate editor and the anonymous reviewers for their valuable comments, which have enhanced the technical quality and lucidity of the paper.

APPENDIX A

PROOF OF PROPOSITION 2.1

Optimizing (13) by solving the corresponding Euler-Lagrange equation for an extremum condition [17] yields the following optimal power allocation [7],

$$I_1^{opt}(s) = \begin{cases} P & s < s_0 \\ \frac{1-e^{-(s_{th}-s)}}{s^2} - \frac{1}{s} & s_0 \leq s \leq s_1 \\ 0 & s > s_1 \end{cases} \tag{A.1}$$

where $s_1 = 1 - W\left(e^{-(s_{th}-1)}\right)$, where again $W(x)$ is the Lambert W-function, and s_0 is determined by requiring $I_1^{opt}(s_0) = P$. Hence, the maximal average rate is obtained by substituting (A.1) into (13), which simplified into (15).

Similarly, for $R_{avg,2}$ (3), when $\nu > s_{th}$, the broadcasting average rate is generally expressed as in (14). Optimizing (14) by solving the corresponding Euler-Lagrange equation for an extremum condition yields the following optimal power allocation, which strongly depends on the selection of s_{th}. For $s_{th} < s_0$,

$$I_2^{opt}(s|s_{th} < s_0) = \begin{cases} P & s < s_0 \\ \frac{1}{s^2} - \frac{1}{s} & s_0 \leq s \leq s_1 \\ 0 & s > s_1 \end{cases} \tag{A.2}$$

where $s_0 = \frac{2}{1+\sqrt{1+4P}}$, and $s_1 = 1$. Note that in this case ($s_{th} < s_0$) the broadcast approach matches the broadcasting power allocation with no side information. The side information here is relevant only for the case the transmitter is informed that $\nu \leq s_{th}$. For $s_0 \leq s_{th} \leq s_1$ we get,

$$I_2^{opt}(s|s_0 \leq s_{th} \leq s_1) = \begin{cases} P & s < s_{th} \\ \frac{1}{s^2} - \frac{1}{s} & s_{th} < s \leq s_1 \\ 0 & s > s_1 \end{cases} \tag{A.3}$$

which means that for $\nu = s_{th}$ a non-infinitesimal power is allocated. That is, for $\nu = s_{th}$ the power allocation is $P - \frac{1}{s_{th}^2} + \frac{1}{s_{th}}$. That is, $\rho_2^{opt}(s|s_0 \leq s_{th} \leq s_1) = \left(P - \frac{1}{s_{th}^2} + \frac{1}{s_{th}}\right)\delta(s - s_{th}) + (\frac{2}{s^3} - \frac{1}{s^2})\mathbf{1}(s_{th} < s \leq s_1)$, where $\delta(x)$ is the Dirac delta function, and $\mathbf{1}(x)$ is the indicator function. For $s_{th} > s_1$, the optimal strategy is to perform single level coding, that is

$$I_2^{opt}(s|s_{th} > s_1) = \begin{cases} P & s < s_{th} \\ 0 & s \geq s_{th} \end{cases} \tag{A.4}$$

Summarizing, the maximal average rate, resulting from the optimal power allocation (A.2)-(A.4) as function of relative value of s_{th}, is given by (16). ■

APPENDIX B

PROOF OF PROPOSITION 4.1

The first step is to show that for the highest (n^{th}) layer Q_n satisfies properties 1) and 2). Then it will be shown that necessarily the same result for holds for Q_{n-1}. For Q_n, it is assumed that all previous $n-1$ layers were successfully decoded and canceled from the receive signal. Then, the outage probability minimization can be written as

$$P_{o,n} = \min_{Q_n,\ \textbf{s.t.}\ \text{tr}(Q_n)=1} \text{Pr}\left(\log\left(1+\alpha_n \mathbf{h}_s Q_n \mathbf{h}_s^{\dagger}\right) < R_n\right). \tag{B.1}$$

where $\alpha_i \in [0,1],\ \forall i = 1,..,n$, and $\sum_{i=1}^{n} \alpha_i = 1$. The minimization of $P_{o,n}$ in (B.1) is identical to the outage case, since there **is** only one last layer, without residual interference. Thus, the covariance matrix Q_n, which minimizes $P_{o,n}$, satisfies properties 1) and 2) [12]. The outage probability minimization for layer $n-1$ is given by

$$P_{o,n-1} = \min_{Q_{n-1},\ \textbf{s.t.}\ \text{tr}(Q_{n-1})=1} \text{Pr}\left(\log\left(1+\frac{\alpha_{n-1}\mathbf{h}_s Q_{n-1}\mathbf{h}_s^{\dagger}}{1+\alpha_n \mathbf{h}_s Q_n \mathbf{h}_s^{\dagger}}\right) < R_{n-1}\right), \tag{B.2}$$

which may be expressed in a simplified form as

$$P_{o,n-1} = \min_{Q_{n-1},\ \textbf{s.t.}\ \text{tr}(Q_{n-1})=1} \text{Pr}\left(\mathbf{h}_s \widetilde{Q} \mathbf{h}_s^{\dagger} < g\right), \tag{B.3}$$

where $g = e^{R_{n-1}} - 1$, and $\widetilde{Q} = \alpha_{n-1}Q_{n-1} - \alpha_n(e^{R_{n-1}} - 1)Q_n$. The minimization in (B.3) can be solved similarly to the outage case. Hence we have that $\widetilde{Q}$ satisfies the properties 1) and 2) of the proposition. Since Q_{n-1} is a linear combination of Q_n and $\widetilde{Q}$, i.e. $Q_{n-1} = \frac{1}{\alpha_{n-1}}[\widetilde{Q} + \alpha_n(e^{R_{n-1}} - 1)Q_n]$, it is straightforward to show that Q_{n-1} also meets properties 1), 2). This can be repeated for Q_{n-2}, and for any layer. ■

REFERENCES

[1] A. Moustakas and S. H. Simon, "Outage capacity with two-bit channel feedback for a two-transmit and single receive antenna system," *Proceedings of IEEE Global Telecommunications Conference, 2003. GLOBECOM '03.*, vol. 2, pp. 844–848, Dec. 2003.

[2] X. Liu, H. Yang, W. Guo, and D. Yang, "Capacity of fading channels with quantized channel side information," *IEICE Trans. Commun.*, vol. 89B, no. 2, pp. 590–593, Feb. 2006.

[3] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Transactions on Information Theory*, vol. 43, no. 6, p. 19861992, Nov. 1997.

[4] T. T. Kim and M. Skoglund, "On the expected rate of slowly fading channels with quantized side information," *submitted to IEEE Transactions on Communications*, June 2005.

[5] S. Shamai (Shitz), “A broadcast strategy for the Gaussian slowly fading channel,” *IEEE ISIT’97, Ulm Germany*, p. 150, June 29–July 4 1997.

[6] A. Steiner and S. Shamai (Shitz), “Multi-layer broadcasting in a MIMO channel,” *The 38th Annual Conference on Information Sciences and Systems (CISS’04), Princeton University*, March 17 - 19 2004.

[7] S. Shamai (Shitz) and A. Steiner, “A broadcast approach for a single user slowly fading MIMO channel,” *IEEE Trans. on Info, Theory*, vol. 49, no. 10, pp. 2617–2635, Oct. 2003.

[8] T. Cover, “Broadcast channels,” *IEEE Trans. on Info. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[9] D. Neuhoff and K. Gilbert, “Causal source codes,” *IEEE Transactions on Information Theory*, vol. 28, no. 5, pp. 701–713, Sept. 1982.

[10] E. Visotsky and U. Madhow, “Spacetime transmit precodingwith imperfect feedback,” *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2632–2639, Sept. 2001.

[11] A. Moustakas and S. H. Simon, “Optimizing multiple-input single-output (MISO) communication systems with general gaussian channels: Nontrivial covariance and nonzero mean,” *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2770–2780, Oct. 2003.

[12] Y. Xie, C. N. Georghiades, and A. Arapostathis, “Minimum outage probability transmission with imperfect feedback for MISO fading channels,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 1084–1091, May 2005.

[13] A. Wiesel, Y. C. Eldar, and S. Shamai (Shitz), “Optimization of the MIMO compound capacity,” *submitted to IEEE Transactions on Wireless Communications*, Nov. 2005.

[14] L. Ozarow, S. Shamai (Shitz), and A. Wyner, “Information theoretic considerations for cellular mobile radio,” *IEEE Trans. Veh. Tech.*, vol. 43, no. 2, pp. 359–378, May 1994.

[15] Y. Yao and G. B. Giannakis, “Rate-maximizing power allocation in OFDM based on partial channel knowledge,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 1073–1083, May 2005.

[16] J. Proakis, *Digital Communications*. New-York: McGraw-Hill, 1995.

[17] I. Geldfand and S. Fomin, *Calculus of Variations*. Mineola, New-York: Dover Publications, Inc., 1991.

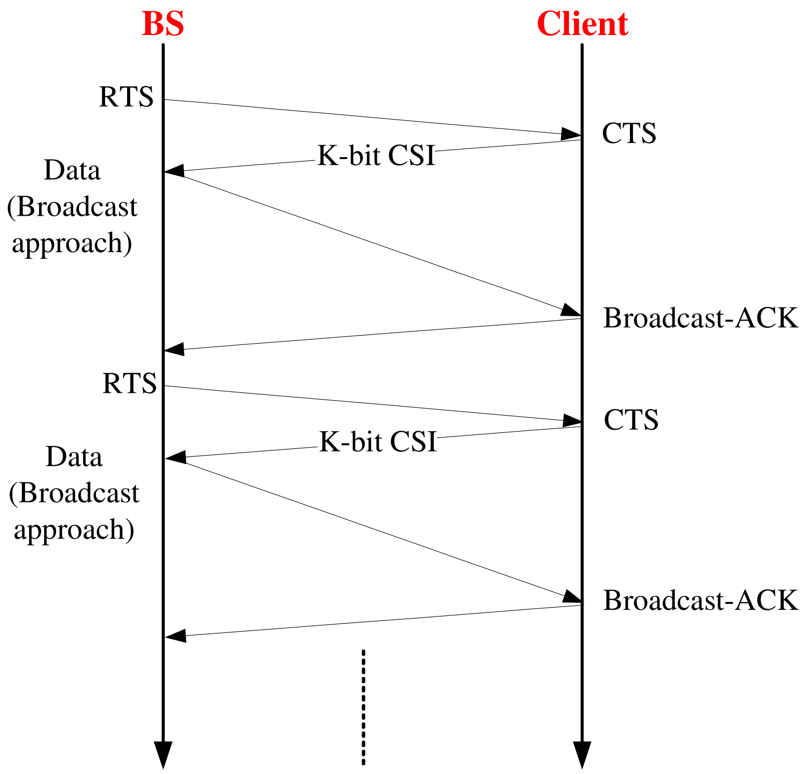

Fig. 1. An exemplary communication protocol for acquiring partial transmit CSI, namely a quantized feedback. Every transmission is initiated with an RTS/CTS sequence, and the CTS includes additional K bits of CSI.

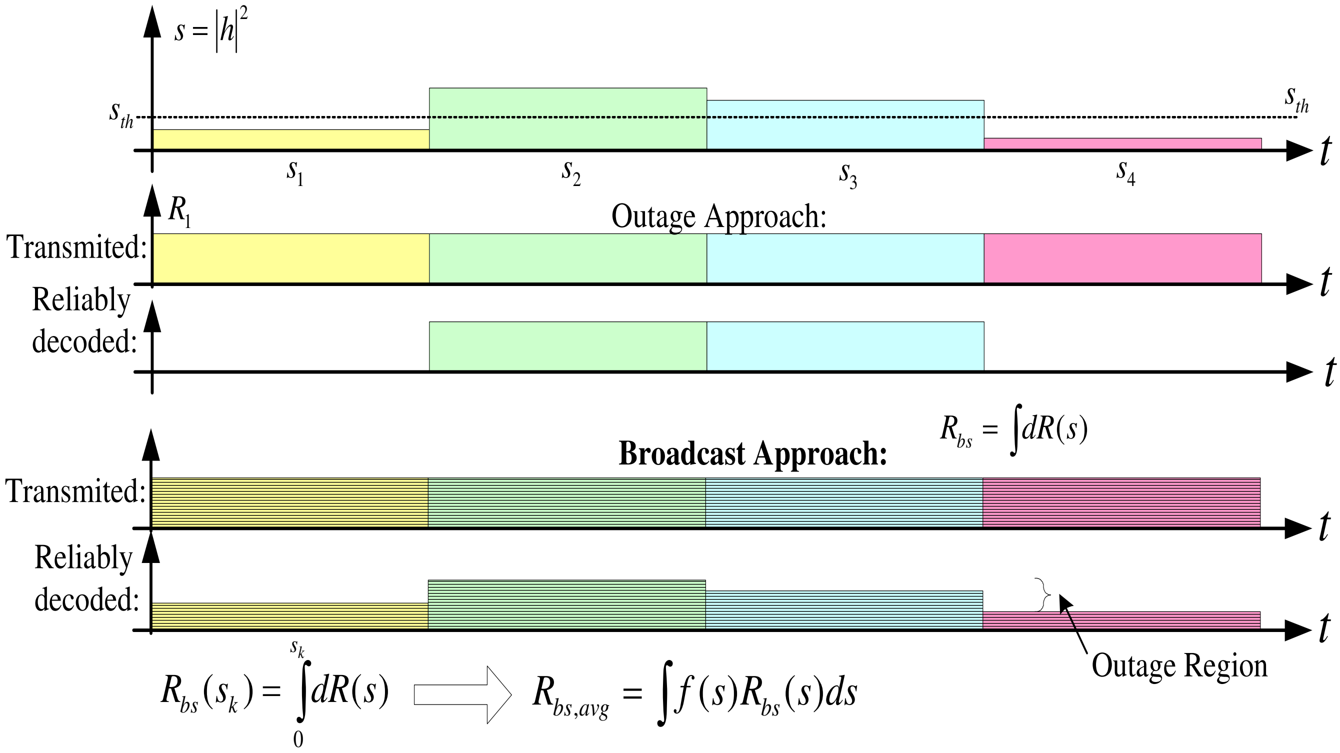

Fig. 2. A timing diagram illustrating the concept of the broadcast approach versus the conventional outage approach.

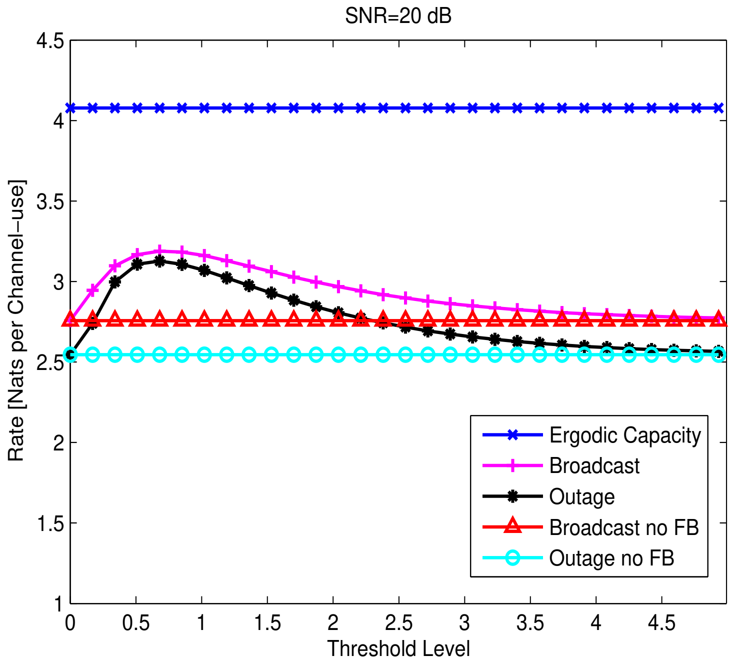

Fig. 3. Broadcasting maximal average rate compared to outage maximal average rate (SNR=20dB). The x-axis represents the feedback information threshold value.

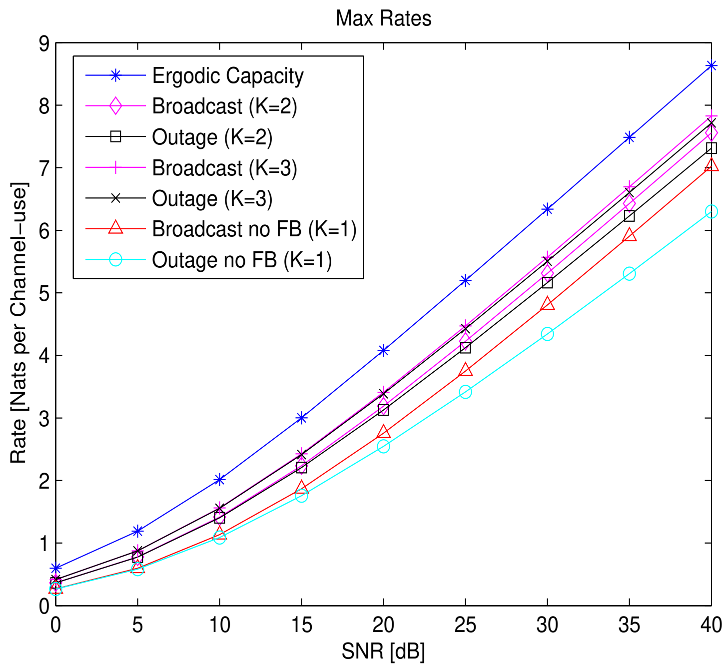

Fig. 4. Broadcasting maximal average rate compared to outage maximal average rate. The case where 1 feedback bit is available is compared to the case where no feedback is available.

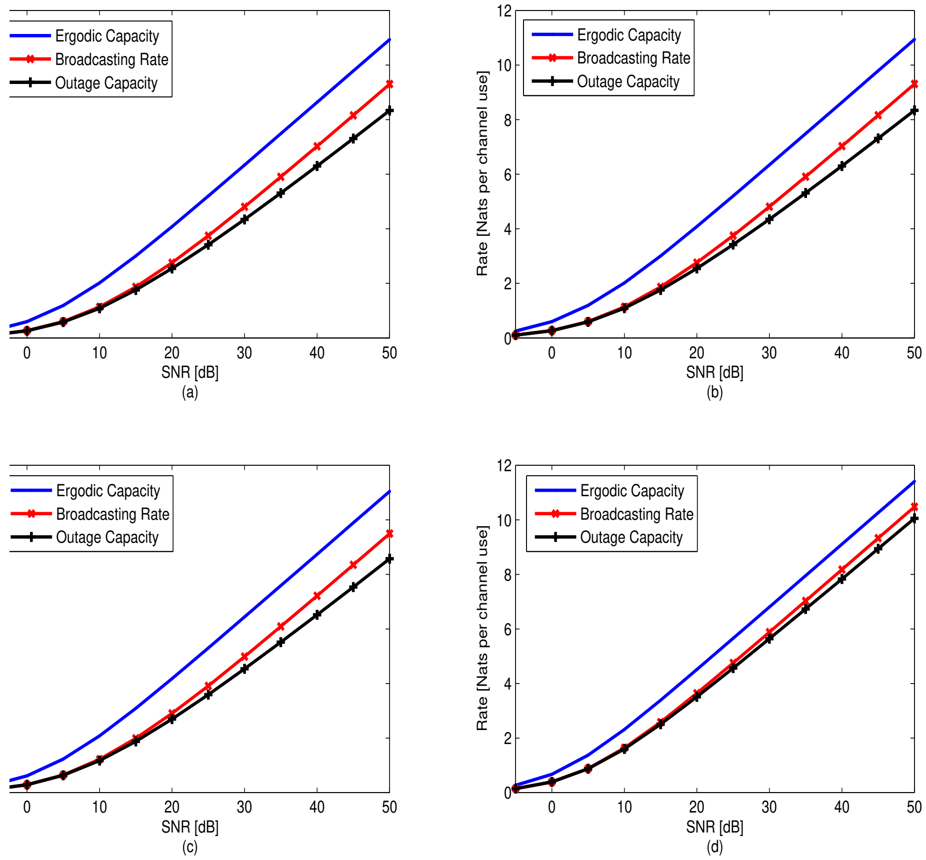

Fig. 5. Maximal achievable average rates of the broadcast approach, outage approach, and the ergodic capacity upper bound. Figure (a) demonstrates the case that $\sigma^2 = 1$, in which case ν_s is Rayleigh distributed. In Figure (b) $\sigma^2 = 0.9$. In Figure (c) $\sigma^2 = 0.5$. In Figure (d) $\sigma^2 = 0.1$.

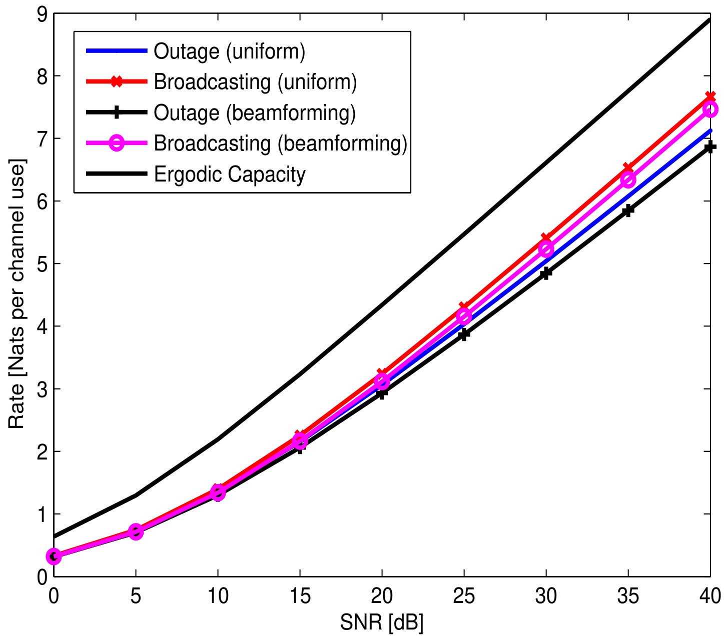

Fig. 6. Maximal achievable average rates of the broadcast approach, outage approach, and the ergodic capacity upper bound. Number of transmit antennas is $N = 2$. Partial CSI is characterized here by $m = 1$, $\sigma^2 = 0.16$.

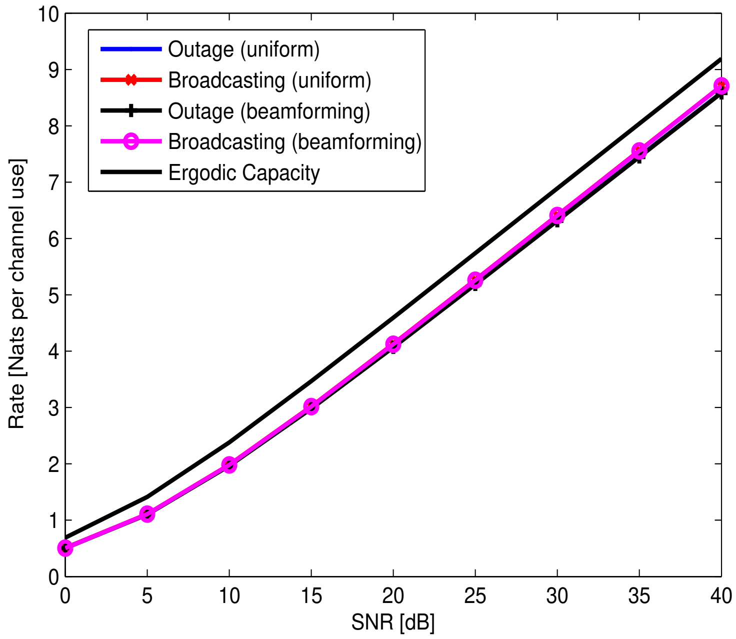

Fig. 7. Maximal achievable average rates of the broadcast approach, outage approach, and the ergodic capacity upper bound. Number of transmit antennas is $N = 2$. Partial CSI is characterized here by $m = 1$, $\sigma^2 = 0.01$.